

デュアル OS 「NINJA」 における擬似 NIC の評価

山本裕馬[†] 榎本圭[‡] 田渕正樹[‡] 伊藤健一[‡] 乃村能成[†] 谷口秀夫[†]

[†]岡山大学工学部 [‡]株式会社 NTT データ

1. はじめに

近年、計算機ハードウェアの性能向上により、複数の計算機環境で実現されてきた多様なサービスを1台の計算機上に集約することが可能になっている。そこで、我々は、Linuxを対象に、1台の計算機上で2つのOSを独立に走行させるデュアル OS 「NINJA」を開発している[1]。「NINJA」では、入出力機器を分割し、各OSに占有させている。

一方、さまざまなサービスへの適用の観点から、一方のOSが占有している入出力機器の機能を他方のOSに提供することが求められている。そこで、NICを対象として、NICとOSの間に擬似ハードウェア(以降、擬似NIC)を実装した。ここでは、擬似NICの基本的な通信性能について報告する。

2. デュアル OS 「NINJA」

2.1 デュアル OS 方式

「NINJA」では、2つのOSを独立に走行させるために、1台の計算機ハードウェアにおける各資源において効果的な共有と占有が必要である。このため、プロセッサのみを共有し、メモリは分割し、各OSに占有させている。また、入出力機器は、OS起動時に指定されたもののみを各OSに占有させている。さらに、タイマ割込みと走行していないOSが占有する入出力機器からの割込みを契機にOSを切替えることによってプロセッサを共有している。

2.2 擬似 NIC

擬似NICは、デュアルOSの利用法を拡張させるために、一方のOSが占有するNICの機能を他方のOSに提供する手法である。

OSはデバイスドライバを用い、I/O命令を発行することによってハードウェアの制御を行う。このため、2つのOSで1つのハードウェアを使用する場合にはI/O命令の監視が必要となる。

デュアル OS 「NINJA」において、擬似NICはNICを占有するOSが持つ。図1に擬似NICの構

成と割込み処理の流れを示し、以降に説明する。

OS1のNICドライバがNICに対してI/O命令を発行しようとする時、OS切替えが行われる。そして、擬似NICがI/O命令の内容を引き継ぎ、I/O命令の種類に応じた処理を行う。次に、図1の(A)~(D)について、処理の内容を説明する。

- (A) 割込みを受ける処理、OS切替えの前処理、およびOSを切替える処理である。
- (B) ドライバで本来行われる処理と、I/O命令発行時にOSを切替える処理である。
- (C) タイマ割込みが発生するまでのOS1での処理とOSを切替える処理である。
- (D) OS1からの復帰処理である。

上記の中で、(A)と(B)のI/O命令、(C)、および(D)で遅延が生じる。なお、(C)についてはNICからの割込みが入るタイミングによって遅延の大きさが異なる。タイマ割込みの場合は周期が10msであるため、遅延時間は0~10msとなる。以降、(A)、(B)のI/O命令、および(D)の処理時間について評価する。

3. 評価

3.1 環境

改造を加えていないOS(以降、オリジナル)とOS1の2つの場合について、UDPによるパケッ

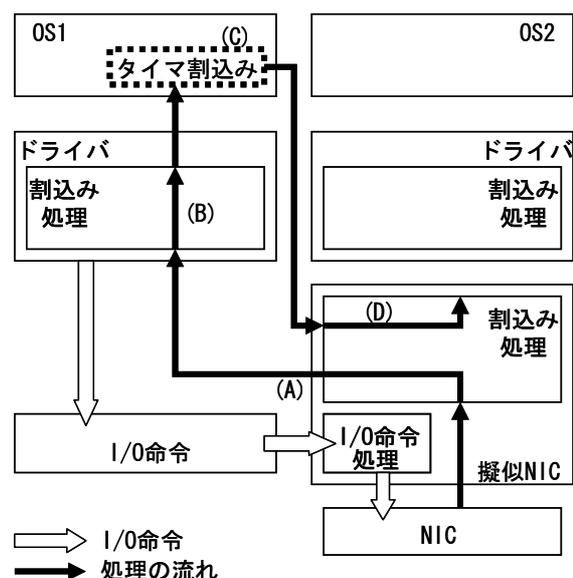


図1 擬似NICの構成と割込み処理の流れ

Evaluation of a virtual NIC on DualOS 「NINJA」
[†]Yu-ma Yamamoto, Yoshinari Nomura, Hideo Taniguchi
[‡]Kei Masumoto, Masaki Tabuchi, Ken-ichi Itoh
[†]Faculty of Engineering, Okayama University
[‡]NTT DATA Co.

トの受信と送信の処理時間を測定した。

測定環境は、CPU : Pentium4 3.0GHz, OS : Linux Kernel 2.4.7, NIC : 3com 3c905-TX である。

測定は CPU のタイムスタンプカウント値を出力する rdtsc 命令を用いた。

3.2 受信

OS1 がパケットを受信する場合の処理の流れは、図 1 に示した割込み処理の流れと同じである。また、オリジナルがパケットを受信する場合は、図 1 の(B)のドライバで行われる本来の処理のみである。

測定結果を図 2 に示す。図 2 は、図 1 の(A), (B), および(D)の処理時間である。

図 2 より、OS1 は、いずれのパケットサイズでも、オリジナルの場合よりも約9.90 μ s 遅くなるのがわかる。なお、遅延時間の詳細を以下に示す。

- 割込みを受ける処理 ; 約0.76 μ s
- OS 切替えの前処理 ; 約1.21 μ s
- OS を切替える処理 ; 約3.99 μ s
- I/O 命令発行時に OS を切替える処理 ; 約3.19 μ s
- OS1 からの復帰処理 ; 約0.75 μ s

I/O 命令は、横取りしない場合に比べて横取りした場合は0.6~0.7 μ s 遅くなった。受信の場合は I/O 命令が5回発行されている。なお、割込みを受ける処理、OS を切替える処理、および OS1 からの復帰処理は、ハードウェアの種類に依存しないため、NIC 以外の擬似ハードウェアを実装した場合も約5.50 μ s の遅延時間が生じると推察できる。

3.3 送信

OS1 がパケットを送信する場合は、受信の場合の処理の流れに加えて、ドライバで送信処理が行われる。

測定結果を図 3 に示す。図 3 は、ドライバでの送信処理、図 1 の(A), (B), および(D)の処理時間である。

OS1 は、いずれのパケットサイズでも、オリジナルの場合よりも約11.98 μ s 遅くなるのがわかる。これは、送信の場合、OS 切替えの前処理がないものの、I/O 命令が10回発行されているためである。

3.4 考察

OS1 は、オリジナルに比べて1パケットあたり受信の場合約9.90 μ s、送信の場合約11.98 μ s 遅くなる。

100Mbps の伝送路で通信を行った場合、1000 バイトのパケットを送信すると、伝送時間は80 μ s になる。この伝送時間に占める擬似 NIC による

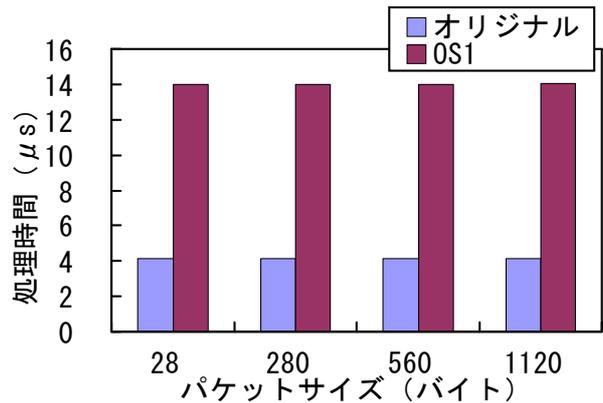


図 2 受信時間

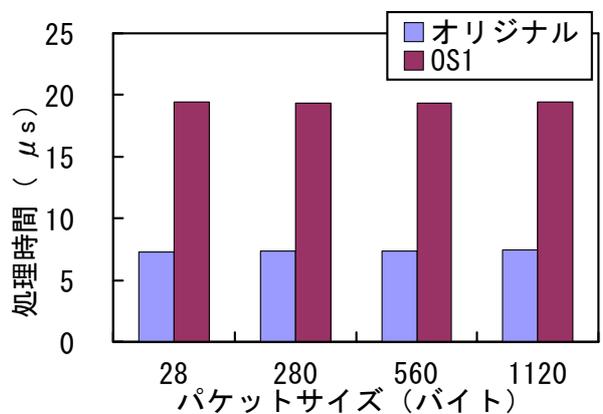


図 3 送信時間

遅延時間の割合は、15%程度になる。なお、パケットサイズが大きくなるほど遅延程度は小さくなる。

また、1つの I/O 命令に生じる遅延は0.6~0.7 μ s であり、ハードウェアの種類に関係なく擬似のために生じる遅延は5.50 μ s であった。これにより、他の PCI ハードウェアに適用した場合に生じる遅延時間を予測することができる。

4. おわりに

「NINJA」における擬似 NIC の性能評価結果を報告した。擬似 NIC によるパケット受信や送信時の遅延は十数マイクロ秒程度であることを示した。今後は、擬似 NIC の詳細な評価を行う予定である。

参考文献

[1]田渕正樹, 伊藤健一, 乃村能成, 谷口秀夫:"二つの Linux を共存走行させる機能の設計と評価",電子情報通信学会論文誌 D(I), 掲載予定